PAPER • OPEN ACCESS

On the features of Hurst Exponent estimates of the Fractional Brownian motion calculated by the R/S-analysis

To cite this article: O Ponomareva *et al* 2021 *IOP Conf. Ser.: Mater. Sci. Eng.* **1047** 012018

On the features of Hurst Exponent estimates of the Fractional Brownian motion calculated by the R/S-analysis

O Ponomareva[1], S Porshnev[1,2] and E Solomakha[1]

[1] The Institute of Radioelectronics and Information Technologies, 620002 Ural Federal University, Mira st. 32, Yekaterinburg, Russia
[2] N.N. Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of SciencesType the author addresses here

E-mail: o.a.ponomareva@urfu.ru

Abstract. The article presents the analysis results of the dependence of the accuracy in estimating Hurst exponent of the Fractional Brownian motion by the R/S-analysis towards the method parameters $L_{\min}$, $L_{\max}$. It is found that the estimation of the Hurst exponent coinciding with its corresponding value is used to generate Fractional Brownian $H^{\langle \mathrm{mod} \rangle}$ motion only when $L_{\max} = L_{\max}^{\langle true \rangle}$. Otherwise, Hurst Exponent Estimate H depending on the value $L_{\max}$ varies in the span [0.25; 1.12]. The result obtained points out that it is necessary to critically revise the results of a number of studies where in order to analyze and forecast the dynamics of complex systems of different nature (for example, in economic ones) the authors employed the R/S-evaluation exponents of the Hurst exponent H of the time series (TS), composed of the exponents characterizing the state of the given system at a certain point.

1. Introduction

Benoit Mandelbrot [1] paved the way for systematic studies of self-similar random processes, called Fractional Brownian motion (FBM), and noted that, apparently, for the first time in an implicit form, FBM had been considered by A.N. Kolmogorov in 1940 [2]. It should be noted that according to [3], the Gaussian random process, one-dimensional Brownian motion, and FBM are defined as:

1) A random variable X is called a Gaussian or a normal one with mean μ and variance σ^2 if it is distributed according to the law:

$$F_X(x) = P(X < x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) d\xi.$$

2) A random process X (t) is called a Gaussian random process if the vector $\{X(t_1), X(t_2), \dots, X(t_n)\}$ has a Gaussian distribution for each finite set of time intervals $\{t_1, t_2, \dots, t_K\}$.

3) A one-dimensional Brownian motion is a Gaussian process X (t) (or a Wiener process) on the interval [a, b] with the following properties:

- X (0) = 0 and the function X (t) is almost always continuous.

- The property of Gaussian increments: a random variable

$$\Delta X = X(t_2) - X(t_1)$$

has a normal distribution $N(\mu,\sigma)$ with zero mean and variance $\sigma^2 (t_2 - t_1)^2$, where $t_2 > t_1$, i.e.

$$P(\Delta X < x) = \frac{1}{\sqrt{2\pi\sigma^2 (t_2 - t_1)}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}\left(\frac{\xi}{\sigma(t_2 - t_1)} \right)^2 \right) d\xi. \quad (1)$$

4) FBM with parameter H is a process X (t) obtaining the following properties.

- X (0) = 0 and the function X (t), is almost always continuous.
- A property of Gaussian increments: random variable

$$\Delta X = X(t_2) - X(t_1)$$

has a normal distribution $N(\mu,\sigma)$ with zero mean and variance $\sigma^2 (t_2 - t_1)^{2H}$, where $t_2 > t_1$, i.e.

$$P(\Delta X < x) = \frac{1}{\sqrt{2\pi}\sigma (t_2 - t_1)^H} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}\left(\frac{\xi}{\sigma(t_2 - t_1)^H} \right)^2 \right) d\xi. \quad (2)$$

It should be noted that:

- at $H = 1/2$ FBM coincides with one-dimensional Brownian motion;
- it follows from (1) that the variance of the increments of the Brownian motion keeps in with the law:

$$D\left[X(t_2) - X(t_1)\right] = E\left[\left(X(t_2) - X(t_1)\right)^2\right] = \sigma^2 |t_2 - t_1|$$

- it follows from (2) that the variance of FBM increments keeps in with the law:

$$D\left[X(t_2) - X(t_1)\right] = E\left[\left(X(t_2) - X(t_1)\right)^2\right] = \sigma^2 |t_2 - t_1|^{2H}$$

It should be noted that that the fractal dimension d, the measure of self-similarity TS, and the Hurst exponent H are related to each other by the following formula $d = 2 - H$

The concepts of one-dimensional Brownian motion and Fractional Brownian motion described above are used as mathematical models of TS, composed of the values of selected quantitative indicators, whose values are determined (measured) at specified discrete moments in time $\{t_1, t_2, \ldots, t_K\}$, in most cases $t_2 - t_1 = t_3 - t_2 = \ldots = t_K - t_{K-1} = const$. Moreover, a priori, it is assumed that the value of the H exponent is determined by the state of the system generating TS under study.

In effect, the task of calculating the estimate of the mathematical model parameter H (2), called the Hurst exponent, has to be solved according to the specified TS X_i $i = \overline{1,K}$. In this regard, the following methods are traditionally used: cumulative variance methods [3]; a method of R/S-analysis [4], a method based on the use of the second derivative of TS; a method based on the use of a non-recursive filter, the coefficients of which coincide with the coefficients of the selected wavelet decomposition (for example, the 5-order "Symlets" wavelet); a method based on the analysis of the energy distribution of the wavelet coefficients over the levels of the TS wavelet packet decomposition (three last methods were implemented in a MATLAB package of fbmesi.m function, which is described in the corresponding MATLAB help section).

It is worth noting, that today, most researchers would apply mathematical models (1), (2) to TS of various nature, following the tradition introduced in the monograph by E. Peters [5], formally, without checking the Gaussian property of their increments. Generally speaking, this circumstance makes it possible to cast some doubt on parameter H estimates (H exponent) obtained in this case and their further interpretations. The drawback mentioned can be found both in numerous econometric studies (see, for example, [6]) and in numerous works devoted to the analysis of Internet traffic (see, for example, [7-8] and others).

Another problem in the interpretation of the H exponent estimates is related to the lack of studies on their accuracy, which, a priori, can be expected to depend on the internal parameters of the method used to estimate the Hurst exponent H. In this regard, the study of this issue is quite relevant.

The article presents the study results on the accuracy dependence of estimates of the Hurst exponents H associated with the Fractional Brownian motion by the R/S-analysis, which is most often used in econometrics, towards the used values of the parameters of this method. The choice of FBD trajectories with the model Hurst exponent $H_1^{\langle \mathrm{mod} \rangle} = 0.3$, $H_2^{\langle \mathrm{mod} \rangle} = 0.5$, $H_3^{\langle \mathrm{mod} \rangle} = 0.75$, synthesized in accordance with the Fourier filtering method [3], as test samples, is due to the fact that trajectory No. 1 is an example of the so-called "pink" noise [5], trajectory No. 2 is an example of classical Brownian motion, trajectory No. 3 is an example of “black” noise [5].

2. Algorithm for the R/S analysis

The R/S-analysis used to estimate a self-similarity index of a time series (TS), called the Hurst exponent (traditionally denoted by the letter H), was proposed in 1951 by the hydrogeologist Harold Edward Hurst [2], who had been analyzing statistical data of annual flows of the Nile River for more than 40 years. This method is implemented by performing the following sequence of actions.

1. The analyzed TS x_k, $k = \overline{1, K}$ is divided into N contiguous intervals with length L $N = \mathrm{fix}(K / L)$, where $\mathrm{fix}(\cdot)$ is the function that cuts off the fractional part of the number.

2. At each of the time intervals selected in item 1, the average TS values are calculated, the aggregate of which will be further called the ensemble of average values $\{\overline{x}_n\}$, $n = \overline{1, N}$:

$$\overline{x}_n = \frac{1}{L} \sum_{i=1+(n-1)L}^{nL} x_i .$$

3. At each of the time intervals selected in item 1, the mathematical expectation of a fragment of the analyzed TS that falls into the given time interval is calculated:

$$\sigma_n = \left[\sum_{i=1+(n-1)L}^{nL} \left(x_i - \overline{x}_n \right)^2 \right]^{0.5} .$$

4. At each of the time intervals selected in item 1, the cumulative deviations of the corresponding TS fragments are calculated:

$$X_n = \sum_{i=1+(n-1)L}^{nL} \left(x_i - \overline{x}_n \right).$$

5. At each of the time intervals selected in item 1, the value of the span of the cumulative deviation is calculated:

$$\left[R_{\max} \right]_n = \max\left(X_n \right) - \min\left(X_n \right).$$

6. At each of the time intervals selected in item 1, the correlation between cumulative deviations and mathematical expectation values of a given TS fragment is calculated:

$$\left[R/S \right]_n = \left[R_{\max} \right]_n / \sigma_n .$$

IOP Conf. Series: Materials Science and Engineering 1047 (2021) 012018 doi:10.1088/1757-899X/1047/1/012018

7. The average value of the correlation between cumulative deviations and mathematical expectations of the corresponding interval of the TS from the ensemble of values $\{[R/S]_n\}$ is calculated:

$$\overline{[R/S]}_N = \frac{1}{N}\sum_{n=1}^{N}[R/S]_n.$$

8. The length of the interval L is increased as long as $L \le L_{\max}\left(L_{\max} \le K/2\right)$ and steps by paragraphs 1-8 of the R/S analysis algorithm are repeated.

9. Using the least squares method (LS), the slope of the line approximating the dependency is calculated $\overline{[R/S]}_N = f\left(\log_2(N)\right)$.

The software implementation of this algorithm in the m-language of the MATLAB package (RSA.m function) is given in Appendix 1.

3. Analysis of the dependence of the estimation accuracy of Hurst exponent of Fractional Brownian Motion towards the parameters of the R/S-analysis algorithm

Let us take a closer look at the results of calculating the Hurst exponent of the typical trajectories of FBM consisting of the values of 2^{20} coordinates of a Brownian particle, generated in accordance with the Fourier filtering algorithm [3] (see figure 1) for the following model values of the Hurst exponent: $H_1^{\langle \mathrm{mod} \rangle} = 0.3$; $H_2^{\langle \mathrm{mod} \rangle} = 0.5$; $H_3^{\langle \mathrm{mod} \rangle} = 0.7$.

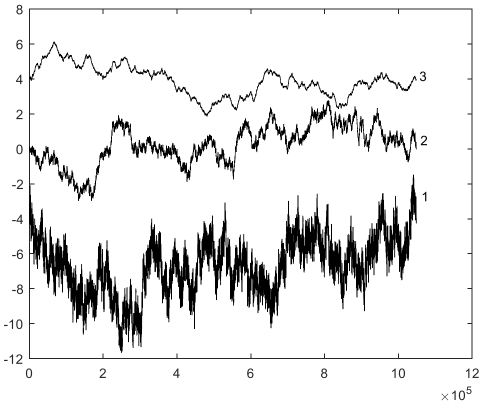

Figure 1. Typical FBM trajectories: 1 – $H_1^{\langle \mathrm{mod} \rangle} = 0.3$; 2 – $H_2^{\langle \mathrm{mod} \rangle} = 0.75$; $H_3^{\langle \mathrm{mod} \rangle} = 0.7$ (trajectories 1, 3 are shifted relative to trajectory No. 2 by 4 units down and up, respectively).

3.1. FBM Trajectory No. 2

A typical graph of dependency $\overline{[R/S]}_N = f\left(\log_2(N)\right)$ as well as intervals $\left[-\mathrm{std}\left([R/S]_n\right), \mathrm{std}\left([R/S]_n\right)\right]$ and the boundaries of the 95% confidence interval of the linear regression of the FBM trajectory No. 2 are shown in figure 2.

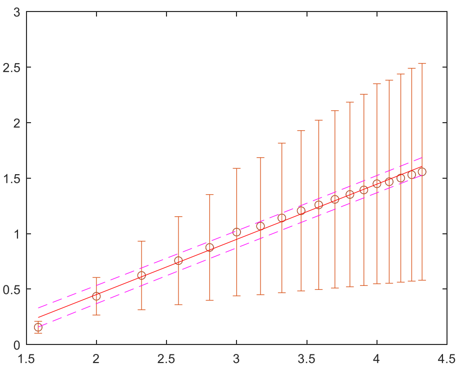

Figure 2. A typical dependency graph $\overline{[R/S]}_N = f\left(\log_2(N)\right)$ (on the example of FBM trajectory No. 2).

Figure 2 shows the interval $\left[-\mathrm{std}\left(\left[R/S\right]_n\right), \mathrm{std}\left(\left[R/S\right]_n\right)\right]$ with an increase in the duration of the intervals of the division of the analyzed TS L (respectively, with a decrease in the number of intervals $N = \mathrm{fix}\left(K/L\right)$) also increases. This result, also, makes it possible to formulate a hypothesis that the closeness in estimation of the Hurst exponent of the given FBM trajectory, generally speaking, will depend on the minimum $(L_{\min})$ and the maximum $(L_{\max})$ values of the L index.

To test this hypothesis, estimates of the Hurst exponent of the discussed FBM trajectory were calculated at various values $L_{\min}$, $L_{\max}$ presented in table 1.

Table 1. Estimates of the Hurst Exponent H2 of FBM Trajectory No. 2, Calculated by the R/S-Analysis.

$L_{\min}$	$L_{\max}$	H	$L_{\min}$	$L_{\max}$	H	$L_{\min}$	$L_{\max}$	H
3	10	0.61317	5	10	0.72267	7	30	0.58261
3	20	0.51566	5	20	0.59180	7	40	0.53756
3	30	0.46779	5	30	0.53065	7	50	0.50560
3	40	0.43713	5	40	0.49246	7	60	0.48208
4	10	0.66953	6	20	0.62528	8	40	0.55766
4	20	0.55553	6	30	0.55771	8	50	0.52402
4	30	0.50092	6	40	0.51601	8	60	0.49862
4	40	0.46641	6	50	0.48660	8	70	0.47861

As it is seen from table 1, in fact, the estimates of the Hurst exponent H2 of FBM trajectory No. 2, calculated by the R/S-analysis, turn out to depend on the values of the parameters $L_{\min}$, $L_{\max}$.

For the convenience of further analysis, the data in table. 1 are given in the form of graphs of dependencies $H = H\left(L_{\max}\right)$ of fixed values $L_{\min}$ (figure 3).

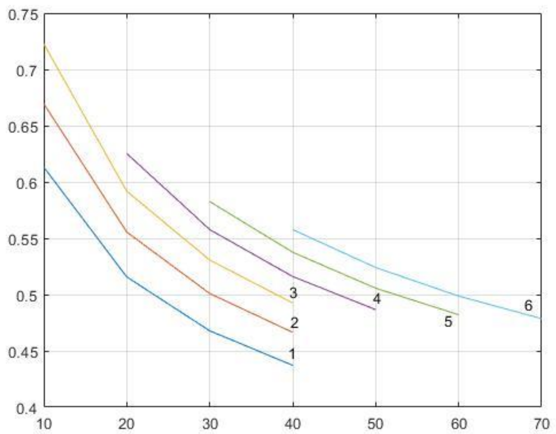

Figure 3. Dependency graphs $H_2 = H_2\left(L_{\max}\right)$: 1 – $L_{\min} = 3$; 2 – $L_{\min} = 4$; 3 – $L_{\min} = 5$; 4 – $L_{\min} = 6$.

Figure 3 illustrates that for each of the dependencies presented in this figure there is a single combination $L_{\min}$, $L_{\max}$ at which the graph of the dependency $H_2 = H_2\left(L_{\max}\right)$ intersects the straight line $H_2 = 0.5$. Due to the discovered dependency feature of the Hurst exponent estimate H_2 of the FBM, obtained using the R/S-analysis, the pairs of the values $L_{\min}$, $L_{\max}$ were calculated at which a difference of the discussed estimate of the Hurst exponent of FBM trajectory No. 2 H_2 of $H_2^{\langle\mathrm{mod}\rangle} = 0.5$ turned out to be the smallest. They are shown as the points $\left(L_{\min_k}, L_{\max_k}\right)$ $k = \overline{1,8}$, $L_{\min_1} < L_{\min_2} < \ldots < L_{\min_8}$, on the plane $\left(L_{\min}, L_{\max}\right)$ (figure 4).

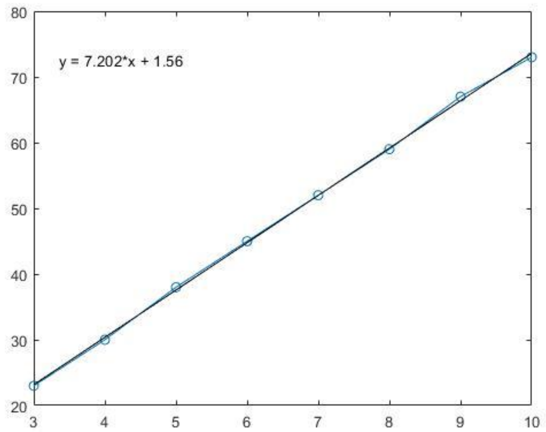

Figure 4. Visualization of pairs of values $\left(L_{\min_k}, L_{\max_k}\right)$ at which a difference of estimating the Hurst exponents of the FBM trajectory No. 2 H_2 of $H_2^{\langle \mathrm{mod} \rangle}$ is the smallest.

Figure 4 shows that the discussed pairs of values are located on a line

$$L_{\max} = 7.202 L_{\min} + 1.560.$$

To find the possible span of estimates for the Hurst exponents of the FBM trajectory No. 2 with an arbitrary choice of pairs of values $\left(L_{\min}, L_{\max}\right)$, the dependencies $H_2 = H_2\left(L_{\max}\right)$ for $L_{\min} \in [3, 24]$, $L_{\max} \in [3; 510]$ shown in figure 5, were calculated.

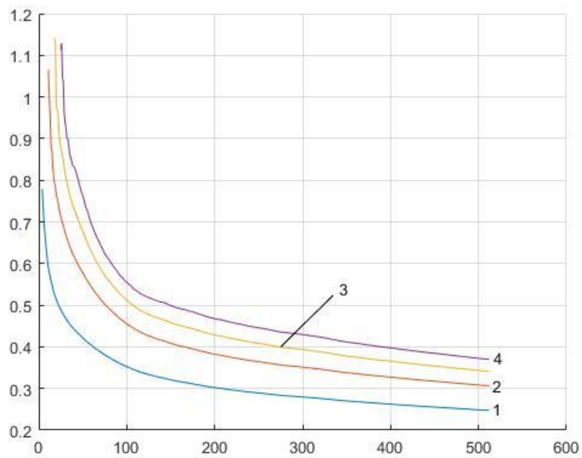

Figure 5. Graphs of dependencies of the estimates of the Hurst exponent of the FBM trajectory No. 2: $H_2 = H_2\left(L_{\max}\right)$: 1– $L_{\min} = 3$, 2 – $L_{\min} = 10$; 3 – $L_{\min} = 17$; 4 – $L_{\min} = 24$.

Figure 5 illustrates that, depending on the correlation between the values $L_{\min}$, $L_{\max}$, the estimates of the Hurst exponent of the FBM H2 will vary in the span [0.25; 1.12].

The pairs of values $L_{\min}$, $L_{\max}$ at which the deviation of the estimate of the Hurst exponent of the trajectory of FBM No. 2 H_2 of $H_2^{\langle \mathrm{mod} \rangle}$ turns out to be the smallest, are presented as the points, $\left(L_{\min_k}, L_{\max_k}\right)$, $k = \overline{1,8}$, $L_{\min_1} < L_{\min_2} < \ldots < L_{\min_8}$, on the plane $\left(L_{\min}, L_{\max}\right)$ (figure 4).

3.2. Trajectories No. 1, 3

Let us consider, for the selected values $L_{\min}$, the dependence of the estimates of the Hurst exponents of trajectory No. 1 $\left(H_1^{\langle\mathrm{mod}\rangle}=0.3\right)$ and trajectory No. 3 $\left(H_3^{\langle\mathrm{mod}\rangle}=0.75\right)$ on the parameter: $L_{\max}$: $H_1=H_1(L_{\max})$, $H_3=H_3(L_{\max})$, respectively, shown in figure 6, 7.

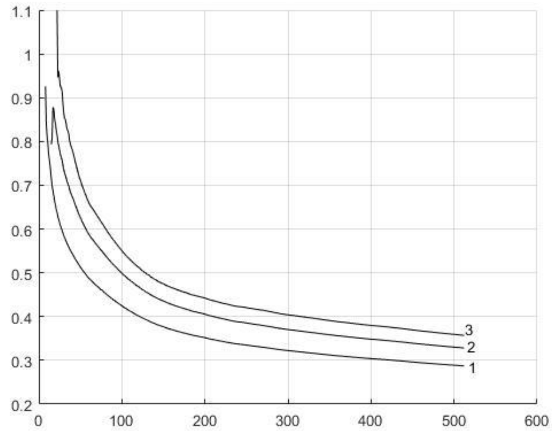

Figure 6. Dependency graphs of the estimates of the Hurst exponents of FBM trajectory No. 1 $H_1=H_1(L_{\max})$: 1 – $L_{\min}=3$, 2 – $L_{\min}=10$; 3 – $L_{\min}=17$; 4 – $L_{\min}=24$.

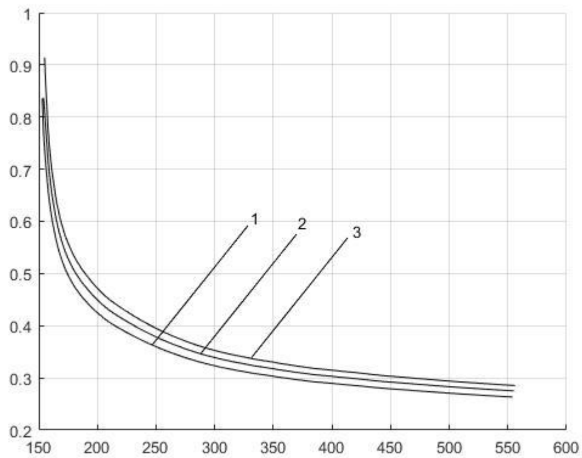

Figure 7. Dependency graphs of the estimates of the Hurst exponents of FBM trajectory No. 3 $H_3=H_3(L_{\max})$: 1 – $L_{\min}=7$, 2 – $L_{\min}=14$; 3 – $L_{\min}=21$.

Figure 6, 7 illustrates that depending on the ratio of the values $L_{\min}$, $L_{\max}$ of the estimates of the FBM Hurst exponents H_1 varies in the span [0.29;1.10], H3 – in the span [0.27,0.91], correspondingly.

The pairs of values $L_{\min}$, $L_{\max}$ at which the deviation of the estimate of the Hurst exponent of the trajectory of FBM No. 1, 2, 3 H_1, H_2, H_3 of the Hurst exponent values used to generate FBM ($H_1^{\langle\mathrm{mod}\rangle}=0.3,\ H_2^{\langle\mathrm{mod}\rangle}=0.5,\ H_3^{\langle\mathrm{mod}\rangle}=0.75$, correspondingly), turns out to be the smallest, are presented as the points, $\left(L_{\min_k}, L_{\max_k}\right)$, $k=\overline{1,8}$, $L_{\min_1}<L_{\min_2}<\ldots<L_{\min_8}$, on the plane $(L_{\min}, L_{\max})$ (figure 8).

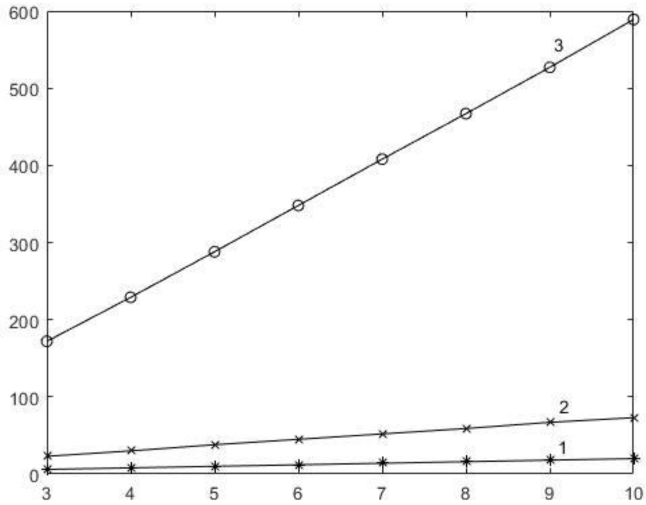

Figure 8. Visualization of pairs of values $\left(L_{\min_k}, L_{\max_k}\right)$ at which a difference of estimating the Hurst exponents of the FBM trajectories No. 1, 2, 3 H_1, H_2, H_3 of the Hurst exponent values used to generate FBM turns out to be the smallest: 1 – $H_3^{\langle \mathrm{mod} \rangle} = 0.75$, 2 – $H_2^{\langle \mathrm{mod} \rangle} = 0.5$, 3 – $H_1^{\langle \mathrm{mod} \rangle} = 0.3$.

Figure 8 shows that at $H_3^{\langle \mathrm{mod} \rangle} = 0.75$ the discussed parameters are found on the line

$$L_{\max} = 2.0 L_{\min},$$

at $H_1^{\langle \mathrm{mod} \rangle} = 0.3$ on the line

$$L_{\max} = 56.6 L_{\min} - 8.87.$$

We suggest that this result can be explained by the fact that the degree of "noisiness" of the FBM trajectory No. 1 ($H_1^{\langle \mathrm{mod} \rangle} = 0.3$) appears to be the highest, the influence of which on the estimate of the Hurst exponent can be reduced by using larger values $L_{\max}$ (in comparison with cases $H_2^{\langle \mathrm{mod} \rangle} = 0.5$, $H_3^{\langle \mathrm{mod} \rangle} = 0.75$).

Thus, the closeness in estimating the Hurst exponent of the Fractional Brownian motion by the R/S analysis appears to be critically dependent on the correlation between the parameters of this method $L_{\min}$, $L_{\max}$.

4. Conclusion

The results of this study confirm, a priori, the hypothesis that the values of the estimates of the Hurst exponent H of the FBM trajectories, calculated using the R/S analysis, will depend on the parameters of this method $L_{\min}$, $L_{\max}$.. It turns out that the estimates of the Hurst exponent H of the FBM trajectories are close to the corresponding values of Hurst exponent $H^{\langle \mathrm{mod} \rangle}$, used to generate FBM trajectories only when specific pairs of parameters values $L_{\min}$, $L_{\max}$. are chosen.

The discussed pairs of parameter values $\left(L_{\min}, L_{\max}\right)$ on the plane are located along straight lines, the slopes of which are determined by the value of the Hurst exponent used to generate the FBM trajectory. The use of pairs of parameter values $L_{\min}$, $L_{\max}$ that do not belong to these straight lines leads to the fact that the values of the estimates values of the Hurst exponent, calculated by the R / S-analysis, will vary in the span [0.25; 1.12].

The described fact allows us to hypothesize that the presence of fundamental limitations of the R / S-analysis method used to estimate the Hurst exponent of real TS. Aт essential condition for finding the maximum close estimate of the fractal dimension of TS to its real value is the choice of a pair of parameter values $L_{\min}$, $L_{\max}$, belonging to the only one straight line on the plane $\left(L_{\min}, L_{\max}\right)$. However, to select the corresponding line, in turn, it is necessary to know the exact value of the fractal dimension of the analyzed TS, which must be determined.

In this case, as the analysis of numerous works shows, in which the R / S-analysis was used to estimate the Hurst exponent of real TS, their authors randomly chose specific parameters $L_{\min}$, $L_{\max}$ and received estimates of the Hurst exponents, on the basis of which the dynamics of complex systems that generated analyzed TS was carried out. However, the probability of a random choice of a pair of parameters $L_{\min}$, $L_{\max}$, corresponding to the condition of maximum proximity of the estimate of the Hurst exponent of the analyzed TS to its actual value is obviously extremely small. In this case, the results of these works (first of all, works on the theory of fractal markets) require their critical analysis, and, first of all, a comparison of estimates of the Hurst exponents calculated by the R / S-analysis method with other available methods which are free from the mentioned drawback, for example, by the accumulated variance method.

Acknowledgement

The work was supported by Act 211 Government of the Russian Federation, contract № 02.A03.21.0006.

References

[1] Mandelbront B and Van J 1968 Ness Fractional Brownian Motions Fractional Noise and Applications *SIAM Review* **10(4)** 422-37
[2] Kolmogorov A 1940 Spiral' Vinera i nekotorye drugie interesnye krivye v gil'bertovom prostranstve *Doklady Akademii nauk SSSR* **26(1)** 115-8 (in Russian)
[3] Crownover R 1995 *Introduction to Fractals and Chaos* (Jones & Barlett Publishers) p 306
[4] Hurst E 1951 Long-term storage capacity of reservoirs *Transactions of the American Society of Civil Engineers* **116(1)** 770-99
[5] Peters E 1996 *Chaos and Order in the Capital Markets: A New View of Cycles Prices and Market Volatility* (John Wiley & Sons) p 288
[6] Zinenko A 2012 R/S-analiz na fondovom rynke *Biznes-informatika* **3(21)** 24-30 (in Russian)
[7] Sheluhin O, Sheluhin O I, Osin A V and Smol'skij S M 2008 *Samopodobie i fraktaly: telekommunikacionnye prilozheniya* (M: Fizmatlit) p 368 (in Russian)
[8] Perepelina A, Popova E, Yangishieva A and Lenshova E 2005 Predprognoznyj analiz obema stoka gornyh rek, kak element ekonomicheskoj bezopasnosti regiona *Vestnik VGU, Seriya: ekonomika i upravlenie* **1** 168-76 (in Russian)